2019

POPULATION GENETIC STRUCTURE OF *SARDINELLA AURITA* AND *SARDINELLA MADURENSIS* IN THE EASTERN CENTRAL ATLANTIC REGION (CECAF) IN WEST AFRICA

Evelyn Takyi
University of Rhode Island, evelyn-takyi@my.uri.edu

POPULATION GENETIC STRUCTURE OF *SARDINELLA AURITA* AND *SARDINELLA MADURENSIS* IN THE EASTERN CENTRAL ATLANTIC REGION (CECAF) IN WEST AFRICA

BY

EVELYN TAKYI

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

IN

BIOLOGICAL SCIENCES

UNIVERSITY OF RHODE ISLAND

2019

MASTER OF SCIENCE THESIS

OF

EVELYN TAKYI

APPROVED:

Thesis Committee:

Major Professor Marta Gomez-Chiarri

Brian Crawford

Rachel Schwartz

Nasser H. Zawia

UNIVERSITY OF RHODE ISLAND

2019

ABSTRACT

Marine fishes represent a valuable resource for the global economy and food consumption, but many species experience high levels of exploitation necessitating effective management plans. Long term sustainability of these resources may be jeopardized from insufficient knowledge about intra-specific population structure. Restriction-site associated DNA sequencing (RAD-seq) methods are revolutionizing the field of population genomics in non-model organisms as they can generate a high number of single nucleotide polymorphisms (SNPs) even when no reference genomic information is available. *Sardinella aurita* and *Sardinella maderensis* are non-model species lacking reference genomes. Using double-digest restriction-site associated DNA (ddRAD) sequencing, we surveyed variation in 6,078 single nucleotide polymorphisms (SNPs) loci identified in *S. aurita* from Mauritania, Senegal, Ghana, Togo, and Benin in West Africa and 6,767 SNPs loci identified in *S. maderensis* from Mauritania, Senegal, Guinea, Togo, and Benin.

Sardinella aurita populations revealed low levels of genetic differentiation (overall F_{ST} value = 0.001 and pairwise F_{ST} values ranging from -0.002 to 0.005, $p > 0.05$) and lack of population structure at the geographic locations surveyed, suggesting the presence of a single panmictic population in this region. Analysis of *S. maderensis* samples also demonstrated that genetic differentiation does not exists across the locations studied (overall F_{ST} = 0.002, and pairwise F_{ST} values ranged from -0.005 to 0.016). More research needs to be performed to extend the geographical and temporal sampling. Overall, results from this research will contribute to our understanding of the distribution of the two

species to assist in the management of the fish resources in the Eastern Central Atlantic region (CECAF) of the Food and Agriculture Organization (FAO).

ACKNOWLEDGMENTS

I would like to thank God Almighty for giving me this opportunity. I am so much grateful and will always be to my major advisor Dr. Marta Gomez-Chiarri for accepting, mentoring and providing any form of support in this program for my success, I say God bless you so much!!! I would also like to thank my committee members Dr. Rachel Schwartz, Dr. Brian Crawford and my defense chair Dr. Jonathan Puritz for their immense support and advice for the data analysis and USAID-Sustainable Fisheries Management Project (SFMP) in Ghana for this scholarship. I will like to also thank Mr. Najih Lazar for his immense contribution to this work. I would like to thank my lab mates: Rebecca Stevick, Erin Roberts, Tejashree Modak, Melisa Hoffman, and Sam Hughes for their warm heart and making the lab an awesome place to be.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 CHAPTER 1: LITERATURE REVIEW

The Eastern Central Atlantic (CECAF) region of the Food and Agriculture Organization (FAO), a management structure for ensuring sustainable use of marine resources, is characterized by the importance of small pelagic resources, including round sardinella (*Sardinella aurita*) and flat sardinella (*Sardinella maderensis*). The two species are strategic products for African populations by providing support for both artisanal and industrial fisheries (FAO, 1999). Their exploitation has great economic importance since they constitute the bulk of the catch in the countries located in the CECAF region by providing the raw material for industries and supporting a large number of jobs such as capture, processing and trade activities in the coastal countries in Africa (Tacon, 2004). The CECAF region, covering the area of West Africa from Morocco to Angola is further divided into two zones; northern and southern. The northern zone includes coastal countries from Morocco to Guinea. The southern zone extends from Sierra Leone to Angola (Figure 2.1).

These two sardinella species are among the most abundant commercially important migratory small pelagic species in West Africa and belong to the Clupedia family. Round sardinellas, *S. aurita,* are found in the Eastern Atlantic from the African coast to Gibraltar and Mediterranean Sea, and in the Western Atlantic from Cape Cod to Argentina. Flat sardinellas, *S. maderensis,* occur in the Mediterranean Sea and in the Eastern Atlantic from Gibraltar southward to Angola. The two species, which are distinguished by several morphological characteristics, are found together in large schools and are mostly combined in statistics and managed together as "sardinella species". As a result of their migratory nature, the degree of mixing among populations from various countries through the

geographic range is crucial to defining fishery management units. As in other pelagic species, eggs and larvae of these species are passively transported by ocean currents, which may carry them to plankton rich coastal areas representing favorable nurseries until they reach the size at which they join the main adult fish stocks. Both species are tolerant to low salinities in estuaries. *Sardinella aurita* prefer clear water with a temperature below 24ºC, while *S. maderensis* prefer warmer waters above 24^0C. Their habitat ranges from near surface, down to 350 m at the edge of continental shelf (Brainerd, 1991).

The abundance of *S. aurita* in most parts of the world is controlled by water temperature and other hydrographic parameters (Failler, 2014). The migration of this species in the northern zone of West Africa shows fish moving between Senegal and Morocco (Boely and Fre`on, 1979). Round sardinellas are found in Senegalese waters in December-January, concentrated along the edges of the shelf between the Cape Verde Peninsula in Senegal and Guinea Bissau until April. Their reproduction is mainly in the continental shelf during the upwelling season but with a major peak identified in May and June (Freon, 1988). At the end of June, the hydrographic conditions of Senegal waters become less favorable, and adults of the species move towards the north to reproduce in the waters surrounding the Arguin Bank in Mauritania from July to August (Boely et al., 1982). In August/September, upwelling is reduced and the fish leave Mauritania and migrate towards the north into Moroccan waters. In the Gulf of Guinea (southern CECAF), the migration of this species has been identified as tied up with the upwelling cycle (Brainerd, 1991) and as the upwelling starts in July, spawning is then at its maximum and the stock spreads out of the eastern half of Cote d'lvoire and toward the east as far as Togo

and Benin. *Sardinella maderensis* on the other hand are less migratory than *S. aurita* and their movements have been well identified in the southern area (Congo-Angola region).

The abundance of both species is known to fluctuate greatly on a decadal time scale in the CECAF region (Longhurst and Pauly, 1987) and has been listed also as part of the International Union for Conservation of Nature (IUCN) Red List of Threatened Species. The small-pelagic stocks in the CECAF region over the last decade have undergone significant fluctuations which can be caused by natural variability because changes in the environment can affect their recruitment (Belvèze, 1991; Cury and Roy, 1989; Zizah et al., 2001). For example, the *S. aurita* stock in the western Gulf of Guinea (fisheries sector in Ghana) has recorded consistent decline in terms of output over the years, with the canoe fishery's annual sardinella catch declining to just over 17,000 metric tons in 2012 from 120,000 metric tons a dozen years earlier (Lazar et al., 2018). An accurate fish stock identification based on genetic studies on the population structure is limited and this limitation has contributed to imperfect scientific management of these fisheries within the region.

Stock management of sardinella species

In fishery management, a unit of stock is normally regarded as a group of fish exploited in a specific area or by a specific method (Carvalho and Hauser, 1995). Population structure of fish stocks serves as the basis of effective fisheries management which defines the spatial boundaries of the stock associated with its seasonal migration and long-term stability within a defined genetic makeup. Given the nature of the marine seafood resources in the CECAF region, potential benefits can be derived if efforts are

made to manage and develop these fishery resource efficiently (Brainard, 1991). The status of the sardinella fishery is monitored and evaluated by the regional working group of the Committee for the Eastern Central Atlantic Fisheries (CECAF) of FAO who ensures improved management of small pelagic resources in West Africa by assessing the state of the stocks and ensuring sustainable use of these resources. Based on the assessment, management recommendations are made for sustainability of the stocks. According to Lazar et al. (2018), the FAO/CECAF Working Group has agreed on the existence of four stocks for these two species in the southern CECAF area within the Gulf of Guinea: 1) Northern zone (Guinea-Bissau, Guinea, Sierra Leone and Liberia); 2) Western zone (Côte d'Ivoire, Togo, Ghana and Benin); 3) Central zone (Nigeria and Cameroon); and 4) Southern zone (Gabon, the Democratic Republic of the Congo, the Congo and Angola). This stock differentiation is an assumption based on management needs and has been defined in the absence of information to match the biological boundaries of these two species with management strategies. Setting fisheries management strategies requires an understanding of fish stock boundaries and fish managers need information on the biological differences and genetic processes of discrete local groups of a species (Palumbi, 1996). Genetic assessment can be used to determine the population structure, determine whether individuals have moved among populations recently or in the distant past, suggest the typical size of a population and, thus, the effective reproducing population (Bernatchez and Wilson 1998; Taylor et al., 2001).

The use of population genetics in fisheries stock management

The most obvious forms of assessing fish populations involve counting or measuring individual fish, but another suite of characteristics that can be very informative is their genes. Although numerous complementary techniques exist to define fish stocks, it is now well established that genetic data analyses are essential to better delineate stock structure for sustainable management (Durand et al., 2013). Population genetic studies are used in determining evolutionary processes occurring in a population. Polymorphism in wild populations is affected by a range of evolutionary forces (Hedrick 2005) where gene-flow reflects migration and leads to increased homogeneity among isolated populations, while genetic drift acting within populations leads to increased levels of differentiation among populations as a cause of random events between generations. Natal homing of individuals and spawning aggregations also contribute to stock structure in populations (Svedang et al., 2007). Inferring the degree of genetic exchange between populations of marine fish species is key to successfully managing exploited populations. This enables the identification of conservation units and assignment of individuals to geographic regions (Dichmont et al. 2012; Funk et al. 2012). Many exploited marine fish are characterized by little intraspecific genetic structuring even over large geographical distances (Bradbury et al. 2008; Ward et al. 1994). Studies have provided valuable information on spatial population structure for aquatic species of management and conservation concern since genetic assessment can be used to identify cryptic species, determine whether individuals have moved among populations recently or in the distant past, and evaluate bottlenecks and founder effects (Beaumont 1994; Nielson 1995).

Assessing these and other properties relies on identifying sets of genetic markers (White et al., 2005; DeHaan et al., 2006). Common genetic markers used in population identification include mitochondrial DNA sequencing, microsatellites, fragment length polymorphisms (RFLPs and AFLPs), single nucleotide polymorphisms (SNPs), and insertion–deletion polymorphisms (indels). The traditional process of marker development is costly (in time and research funding) and usually results in the generation of very few working markers. A decrease in cost of sequencing has allowed for the development of new high-throughput technologies for genotyping and population genomic studies. The genotyping-by-sequencing approach used for instance in restriction site associated DNA sequencing (RAD-Seq) combines the power of high throughput sequencing and large-scale polymorphism genotyping in one step (Baird et al. 2008; Hohenlohe et al. 2010).

Past Application of Genetic Techniques in Sardinella Species from West Africa

Population genetic studies have been carried out in other sardinella species such as *S. albella* in the Persian Gulf and Sea of Oman using mitochondrial DNA (Rahimi, S, P, Sh, & Rahnema, 2016), *S. longiceps* in Indian Ocean waters using microsatellites (Sukumaran et al., 2017), and *S. aurita* in the coastal waters of Florida, USA, using protein electrophoresis (Kinsey et al., 1994). However, population genetic studies dealing with species of economic importance along the rest of the West African coastline are rare, mostly using weakly polymorphic allozyme and maternally inherited mitochondrial DNA (mtDNA) (Durand et al., 2013). Atarhouch et al. (2006) demonstrated that a local Moroccan population of sardine (*Sardina pilchardus*) was genetically depleted as a probable cause of intensive fishing in the recent past. (Chi, 1998) studied allozyme

variability in *S. aurita* from the Congo, Ghana, Ivory Coast and Venezuela. They recorded three polymorphic loci but extremely low levels of diversity and low differentiation between populations. More research is needed to identify the spatial structure of the round sardinella (*Sardinella aurita*) and flat sardinella (*Sardinella maderensis*) stocks within the CECAF FAO region.

The Use of RAD-seq Technique in Population Genetic Studies

The resolution of genetic differentiation allowed by a large number of genome wide polymorphic markers should enhance inference of neutral population structure (Benestan et al. 2015; Lamichhaney et al. 2012; Narum et al. 2013). One of the most popular high-throughput genotyping methods currently available, RAD-seq, combines restriction enzyme digestion of the genome with high throughput sequencing. RAD-seq is particularly relevant for non-model organisms as it allows discovering and genotyping thousands of single nucleotide polymorphisms (SNPs) in hundreds of individuals rapidly and at low cost regardless of size of the genome and prior genomic knowledge (Baird et al. 2008; Puritz et al. 2014). This methodology enables not only genotyping and SNP discovery, but also more complex analyses such as quantitative genetic, phylogeographic studies, and population differentiation studies. Examples of studies in marine fishes that have used RAD-Seq include genetic marker discovery in threespine stickleback (Catchen et al., 2014), a study of the neutral structure in populations of Pacific lamprey (Hess et al. 2013), and resolution of fine-scale structure in Atlantic salmon (Bourret et al. 2013). Consequently, a variety of RAD-seq protocols are increasingly used to identify and genotype genome-wide markers in non-model marine species to directly inform conservation and management efforts (e.g.

Corander et al. 2013; Hess et al. 2013; Larson et al. 2014; Puebla et al. 2014, Puritz et al. 2016).

2 CHAPTER 2

POPULATION GENETIC STRUCTURE OF *SARDINELLA AURITA* AND *SARDINELLA MADERENSIS* IN THE EASTERN CENTRAL ATLANTIC REGION (CECAF) IN WEST AFRICA

By

[1]Evelyn Takyi; [2]Jonathan B. Puritz, [3]Najih Lazar, [4]Marta Gomez-Chiarri

To be submitted to the Journal of Coastal Zone Management

[1]Master' student, Department of Cell and Molecular Biology, University of Rhode Island, Kingston, RI 02881. Email: evelyntakyi72@yahoo.com

[2]Department of Biological Sciences, University of Rhode Island, Kingston, RI 02881,

[3]Coastal Resources Center, Graduate School of Oceanography, University of Rhode Island,

[4]Department of Fisheries, Animal & Veterinary Science, University of Rhode Island, Kingston, RI 02881. Email:gomezchi@uri.edu

ABSTRACT

Marine fishes represent a valuable resource for the global economy and food consumption, but many species experience high levels of exploitation necessitating effective management plans. Long term sustainability of these resources may be jeopardized from insufficient knowledge about intra-specific population structure. Restriction-site associated DNA sequencing (RAD-seq) methods are revolutionizing the field of population genomics in non-model organisms as they can generate a high number of single nucleotide polymorphisms (SNPs) even when no reference genomic information is available. *Sardinella aurita* and *Sardinella maderensis* are non-model species lacking reference genomes. Using double-digest restriction-site associated DNA (ddRAD) sequencing, we surveyed variation in 6,078 single nucleotide polymorphisms (SNPs) loci identified in in *S. aurita* from Mauritania, Senegal, Ghana, Togo, and Benin in West Africa and 6,767 SNPs loci identified in *S. maderensis* from Mauritania, Senegal, Guinea, Togo, and Benin.

Sardinella aurita populations revealed low levels of genetic differentiation (overall F_{ST} value = 0.001 and pairwise F_{ST} values ranging from -0.002 to 0.005, $p > 0.05$) and lack of population structure at the geographic locations surveyed, suggesting the presence of a single panmictic population in this region. Analysis of *S. maderensis* samples also demonstrated that genetic differentiation does not exists across the locations studied (overall F_{ST} = 0.002, and pairwise F_{ST} values ranged from -0.005 to 0.016). More research needs to be performed to extend the geographical and temporal sampling. Overall, results from this research will contribute to our understanding of the distribution of the two

species to assist in the management of the fish resources in the Eastern Central Atlantic region (CECAF) of the Food and Agriculture Organization (FAO).

INTRODUCTION

Fish is a primary source of protein for at least one billion people and can contribute as much as 80% of the animal protein consumed in the world (FAO, 2009b). The commercial value of small pelagic fishes (oil sardines and anchovies) is low in export market, but in developing countries, they contribute to a substantial portion of the income from fishing due to their abundance (Sukumaran et al., 2016). The two sardinella species *Sardinella aurita* and *S. maderensis* are among the most abundant commercially important migratory small pelagic species in West Africa. These species have a wide distribution in the Atlantic and Pacific Oceans and the Mediterranean Sea. They belong to the Clupedia family and occur in the eastern Atlantic from Gibraltar southward to Angola (Bureau & Resources, 1999). The abundance of these species is known to fluctuate greatly on a decadal time scale in the Eastern Central Atlantic region (Longhurst and Pauly 1987) and have also been listed as part of the International Union for Conservation of Nature (IUCN) Red List of Threatened Species. The monitoring of the status of the small pelagic stocks is determined by the regional working group of the Committee for the Eastern Central Atlantic Fisheries (CECAF) of the Food and Agriculture Organization (FAO) (Lazar et al., 2018). For management purposes, the FAO/CECAF Working Group has agreed on the existence of four stocks for these two species in the southern CECAF region: 1) Northern zone (Guinea-Bissau, Guinea, Sierra Leone and Liberia); 2) Western zone (Côte d'Ivoire, Togo, Ghana and Benin); 3) Central zone (Nigeria and Cameroon); and 4) Southern zone (Gabon, the Democratic Republic of the Congo, the Congo and Angola). In the absence of information to match the biological boundaries of these two species with management strategies, this stock differentiation is an assumption based on management needs.

Fish managers need information on the biological differences between discrete local groups of a species as well as an understanding of the genetic and ecological processes that influence their discreteness. The genetic structure of fish populations is important not only because of fundamental interest in biotic evolution (Tudela et al., 1999) but also for the management of fisheries (Roldan et al., 2000). Even in general marine fishes show lack of genetic differentiation, recent studies using advanced markers have provided evidence for some level of differentiation in many marine fishes as these markers are able to resolve signatures of selection and adaptation in response to environmental and habitat change (Wang et al., 2013, Candy et al., 2015, Brennan et al., 2016). Management strategies should be aimed at conserving intra-specific genetic diversity as it has profound implications in deciding potential for recruitment and population recovery (Teacher et al., 2013). Hence, misdirected management actions without knowledge about the stock structure of marine fishes may result in inability to recover from environmental impacts (Sukumaran, et al., 2017).

An accurate fish stock identification based on genetic studies on the stock structure is limited for marine pelagic species in the CECAF region. This limitation has contributed to imperfect scientific management of fisheries within the region. Previous population genetic studies in sardinellas relied on mostly using weakly polymorphic allozyme and mitochondrial DNA (mtDNA) (Durand et al., 2013). Kinsey et al., (1994) studied the population structure of *S. aurita* in the coastal waters of Florida, USA, using protein electrophoresis, revealing low levels of genetic differentiation and lack of population structure. Chi (1998) studied allozyme variability in *S. aurita* from the Congo, Ghana, Ivory Coast and Venezuela, recording three polymorphic loci but low levels of diversity

and low differentiation between populations. Because of the importance of identification of stock structure in stock assessment (Whithead, 1985), we investigated the genetic population structure of *S. aurita* and *S. maderensis* in five sites located along the Eastern Central Atlantic region using Restriction site-Associated DNA sequencing (RADSeq), a sensitive genotyping tool which has the ability to identify and score thousands of genetic markers (SNPs, indels) randomly distributed across the target genome from a group of individuals using Illumina next generation sequencing technology (Davey & Blaxter, 2010). Information from this study will be useful in the stock management of these important fish species by the CECAF working group for accurate delineation of the stocks of this two sardinellas in the CECAF region.

METHODS

Sample Collection and DNA extraction

Sample collection of *S. aurita* and *S. maderensis* was coordinated by Najih Lazar, fisheries advisor on the USAID-Sustainable Fisheries Management Project (SFMP) in Ghana. A total of fifty (50) specimens each of round sardinella were obtained from landing sites in selected countries in 2016 and 2017 (Figure 2.1, Table 2.1). The dorsal fin from the base was cut and placed in a 1.5-2 ml polypropylene vial. The fin tissue was then preserved in 95% ethanol and stored at -20^{0}C until it was ready to be shipped to the University of Rhode Island for extraction. Genomic DNA from fin clips were extracted using Qiagen DNeasy kit (Qiagen, Valencia, CA) following the manufacturer's protocol. DNA quality and concentration were checked using Nanodrop 2000 and Qubit 2.0 fluorometer instruments. Samples were stored at -20° C until processing.

Table 2.1: Collection sites, dates, and number of collected samples for Sardinella aurita and S. maderensis.

Country	Site coordinates		Collection date	Number of samples collected	
	LONG.	LAT.		*S. aurita*	*S. maderensis*
Mauritania	-15.978	18.0841	2016	50	50
Senegal	-16.85	14.179	2016	50	50
Liberia	-10.797	6.3	2016	0	50
Guinea	-11.352	9.932	2017	50	50
Ghana	0.994	5.898	2016	50	25
Togo	1.301	6.151	2016	50	50
Benin	2.435	6.351	2016	40	50
Nigeria	3.406	6.465	2016	0	50

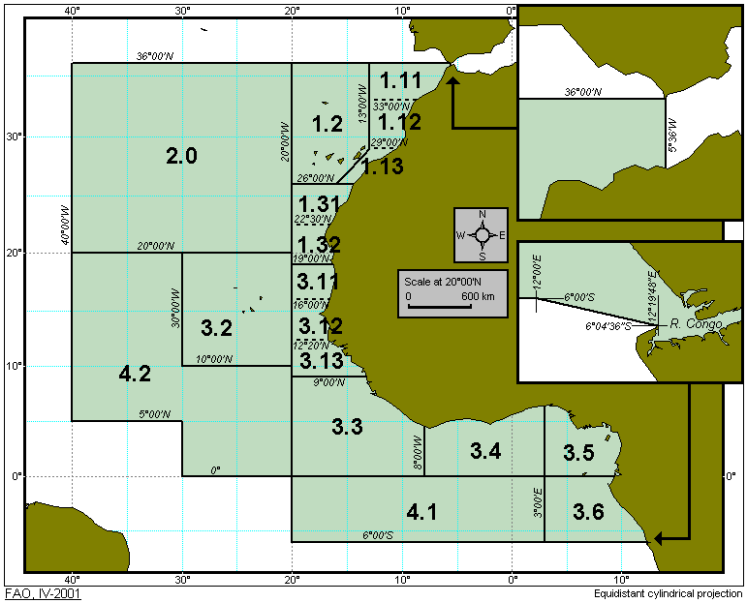

Figure 2.1: Map of areas sampled in West Africa. Areas sampled are from subdivision point 1.32 to 3.4.

Library preparation and sequencing

Library preparation for double digest RAD (ddRAD) and sequencing were carried out by Texas A&M Corpus Christi sequencing Core Center. This protocol was used for both species *S. aurita* and *S. maderensis*. Extracted genomic DNA was normalized to a concentration of 5ng in 50µl in 96 well plates. Restriction enzyme test was performed to determine which restriction enzyme would provide the most cut size combinations and will best target a higher number of loci and coverage. The most consistent enzyme combination across the samples was Hin1II and TasI. Genomic DNA was then processed into ddRAD libraries using these restriction enzymes purchased from New England Biolabs (NEB) following the ddRAD protocol (Peterson et al., 2012). Agarose gels (1%) were run on the DNA to determine the quality of the DNA and samples that had low molecular weight smears were used for SPRIselect size selection using a 0.4X ratio of SPRIselect beads to

DNA volume. Fluorescent quantification of samples was carried out using AccuBlue High Sensitivity solution and a standard curve. An aliquot of 100ng of each sample was transferred to a new working plate and concentrated to 15µl DNA cleanup using 30µl AMPureXP beads (2X bead to DNA ratio), leaving the beads in after elution. Restriction digest was performed using NEB enzymes Hin1II and TasI and DNA was cleaned up using 22.5µl PEG NaCl buffer (1.5X buffer to DNA ratio) with retained beads from first cleanup. Sample concentrations were normalized and adaptor ligated. Samples were pooled with unique barcodes and pools cleaned using AMPureXP beads at a 1.5X ratio of AMPure beads to DNA. Size selection was carried out using Blue Pippin size select targeting a range of 570-645bp. Libraries were PCR amplified using different uniquely indexed PCR primers for each pool. Each pool was separated into eight different PCR reactions. PCR product were pooled and cleaned twice using AMPureXP beads (1X ratio of beads to PCR product for both cleanups) after which pools were run on AATI Fragment Analyzer using the HS NGS kit to get the sizing of the pools. qPCR was performed on pools using KAPA Library Quant kit and combined in equimolar ratios adjusting as necessary for sample number after which it was send to sequencing center and libraries were sequenced in one lane on an Illumina HiSeq4000.

De novo Assembly, Read Mapping, SNP discovery

Quality filtering of raw reads and demultiplexing based on barcode was conducted using process radtags in the Stacks software package (Catchen et al., 2011). The dDocent pipeline (version 2.5.2) was used for read trimming and *de novo* assembly (Puritz et al., 2014). For quality filtering, the program trimmomatic v0.32 (Bolger, Lohse, & Usadel, 2014) was used to trim low quality bases that are below quality score of 20 from the

beginning and end of reads, and an additional sliding 5bp window that will trim bases when the average quality score drops below 10 and removes sequences corresponding to the Illumina adapters.

For *de novo* assembly of each of the reference genomes (*S. aurita* and *S. maderensis)* using the dDocent pipeline, unique sequences in each of the individuals were identified and their coverage counted in the entire data set. Results of the unique sequences with a coverage level of one(1) to twenty(20) were tabulated and unique sequences with a coverage level of four(4) was selected. Next was to select the unique sequences with a coverage of four that appear in most of the individuals for the assembly. A cut off value of the number of individuals a unique sequence should at least appear in was five(5) individuals. Next, the unique sequences with the selected cut off values of four and five were collapsed into FASTA format and assign the header (contig) for the next steps in the *denovo* assembly. The forward reads from each of the contigs were extracted from the FASTA file and the program CD-hit (Fu et al., 2012; Li & Godzik, 2006) was used to cluster the forward reads by 80% similarity into RAD loci. Next, the assembly program Rainbow (Chong, Ruan & Wu, 2012) in the dDocent pipeline was also used to recluster the results of the CD-hit program based on 90% similarity into groups representing alleles at the RAD loci. The longest contig for each cluster was selected as the representative reference sequence for that RAD locus. Finally, clustering of the reference sequences were rechecked by clustering again based on an overall sequence similarity of 90% using the program CD-HIT (Fu et al., 2012; Li & Godzik, 2006) and the reference assembly outputted as a FASTA file.

SNP calling and genotyping were also performed using the dDocent pipeline. For read mapping, reads were mapped to the *denovo* reference file in the FASTA file using the MEM algorithm of BWA (Li & Durbin, 2009; Li & Durbin, 2010) with mismatch parameter lowered from 4 to 3, and the gap opening penalty lowered from 6 to 5. The program Freebayes (Garrison & Marth, 2012) was used to obtain raw variant calls and SNP genotyping which were subjected to several filtering steps to reduce false positives in the SNP calls.

SNP filtering

SNP filtering of the data was performed separately for each species. The criteria that was used in the filtering of the raw variants include the use of VCFtools (Danecek et al., 2011) and custom bash scripts. First, variants that had been successfully genotyped in 50% of individuals with a minimum quality score of 20 and a minor allele count of 3 were retained. Next, loci with a minimum depth less than 5 and minor allele frequency less than 1% were removed. Next, individuals with 60% missing data was filtered using the script (filter_missing_ind.sh) (Puritz et al., 2014). The script (pop_missing_filter.sh) (Puritz et al., 2014) was used to filter out loci by population with 90% missing data. The next filter was based on freebayes generated VCF file using the criteria such as site depth, allelic balance at the heterozygote, properly paired site and maximum mean depth using the dDocent_filters script (Puritz et al., 2014). Variant calls were then decomposed into SNP and INDEL calls with INDELS removed using VCFtools to produce a VCF file of only SNP calls. The bash script rad_haplotyper.pl (Hollenbeck et al., 2017) was used to filter out possible paralogs, possible low-coverage sites, and previously missing genotypes. Loci

with minor allele frequency threshold (maf) < 0.05 were filtered, in order to remove uninformative SNPs. SNPs were then filtered to only include loci with two alleles using VCFtools. The program BAYESCAN (Foll and Gaggiotti, 2008) was used to identify individual outlier loci. The program was run with all default values, with 30 pilot runs and a thinning interval of 100. Significance of outlier loci was determined using a q-value which directly corresponded to a false discovery rate of 0.05. The outlier loci were excluded from SNP loci and the neutral SNPs were utilized for downstream analysis.

Evaluation of genetic diversity between populations

Population genetic statistics (Wright's F statistics F_{IS}, and observed and expected heterozygosity *Ho, He*) were calculated for the populations using divBasic function in diversity package in R (Keenan et al., 2013). The populations program in the stacks software (Catchen et al. 2011) was used to estimate the sites in each population, percentage polymorphic sites and the average frequency of the major allele (P) at the sites. Deviations from Hardy-Weinberg equilibrium was assessed using GENEPOP v4.0 (Rousset 2008). The global estimate for genetic differentiation (F_{ST}) across all samples and loci was calculated following Weir and Cockerham (1984) using the program Adegenet (Jombart, 2008) in R statistical package.

Evaluation of population structure

A UPGMA dendrogram with bootstrap support to visualize the genetic distance between populations was calculated using the function aboot in the poppr package in R with 1000 bootstrap replicates (Jombart, 2008). Node labels represent bootstrap support greater than 50% (>50%). The diffcalf function in the R package diveRsity (Keenan et al.,

2013) was used to calculate the pairwise F_{ST} values for each population and perform significance of genetic differentiation calculation of 95% confidence interval. A higher level of population structure thus individuals nested within populations and population nested within geographic region (southern and northern CECAF) was estimated using an Analysis of Molecular Variance (AMOVA) based approach (Excoffier et al. 1992) implemented in GenoDive (Meirmans and Van Tienderen, 2004). Discriminant analysis of principal components (DAPC) was conducted in R with the package Adegenet 1.3.0 (Jombart 2008). The find.clusters function was used to determine the number of clusters using a graph of Bayesian Information Criterion (BIC). Discriminant Analysis of Principal Component (Dapc) function was used to determine the individual membership probability for each individual. Posterior assignments of each individual were visualized using a composite stacked bar plot known as compoplot which was collapsed into a pie chart showing the assignment of each individual to a group.

RESULTS

Sardinella aurita: de novo assembly and SNP filtering

Samples from six populations of *S. aurita* (Mauritania, Senegal, Guinea, Ghana, Togo and Benin) were sequenced and genotyped using ddRAD sequencing. Only individuals with greater or equal to 500,000 reads were used in the analysis. For *de novo* reference assembly, 82,324 contigs were obtained. There were 115 individuals from the six populations initially used in obtaining a total of 1,256,073 SNP loci. However, after mapping and SNP filtering of this six populations, all samples from the Guinea population (20 samples) were removed from the analysis based on small number of contigs (supplemental table 2.8) mapped to the reference assembly and the high number of individuals (14 out of 20) lost in the filtering process due to a high percentage of missing loci. Removal of the Guinea population led to a total of 95 individuals from five populations being included in the analysis. A total of 1,256,073 variants were obtained using the 95 individuals and subjected to SNP filtering. Three outlier loci were detected and removed from the SNP VCF file retaining 6,078 neutral SNPS loci which passed the quality filtering and 81 individuals out of 95 individuals (14 individuals across all the population with more than 60% missing data were removed from the analysis) for downstream analysis.

Genetic diversity in Sardinella aurita

The average observed (Ho) and expected (He) heterozygosity was observed to be similar across sampling sites, with no significant departure from Hardy-Weinberg Equilibrium (HWE). Observed heterozygosity (Ho) values ranged from 0.31 to 0.32 and expected heterozygosity (He) ranged from 0.26 to 0.27 (Table 2.2). The overall inbreeding

coefficient (F_{IS}) considering all the population was -0.562 and the mean F_{IS} values ranged from -0.42 to -0.44 (considered as zero). The number of polymorphic loci identified in each location ranged from 5,133 in Togo to 5,675 in Mauritania. Nucleotide diversity was similar across all sampling sites and no private alleles were observed for any of the sampled populations.

Table 2.2: Summary genetic statistics for all sampling locations. Population, N (number of samples), total number of sites analyzed, polymorphic sites in each population, average frequency of the major allele (P), He (expected heterozygosity), Ho (observed heterozygosity), F_{IS} (Wright statistics F).

Population	**N**	**Sites**	**Polymorphic loci**	***P***	***He***	**Ho**	***F_{IS}***	**Nucleotide diversity**
Mauritania	16	6,078	5,675	0.802	0.33	0.27	-0.44	0.288
Senegal	17	6,078	5,359	0.804	0.32	0.27	-0.42	0.290
Ghana	17	6,078	5,597	0.806	0.31	0.27	-0.42	0.284
Togo	18	6,078	5,133	0.806	0.33	0.26	-0.44	0.289
Benin	13	6,078	5,335	0.809	0.32	0.26	-0.43	0.286

Genetic Population Structure in Sardinella aurita

To assess genetic differentiation among *S. aurita* collected from locations through the CECAF range, we calculated the global F_{ST} and performed pairwise F_{ST} comparisons between the locations (Table 2.3). No significant differentiation was observed between these populations, with an overall F_{ST} value of 0.001 and pairwise F_{ST} values ranging from -0.002 to 0.005 ($p > 0.05$; Table 2.3).

Table 2.3: Pairwise FST values calculated using 6,078 SNP loci (P > 0.05)

	Mauritania	**Senegal**	**Ghana**	**Togo**
Mauritania				
Senegal	0.003			
Ghana	0.001	0.003		
Togo	0.005	-0.002	0.004	
Benin	0.001	0.002	0.000	0.002

Analysis of Molecular Variance (AMOVA) was used to measure the partition of the genetic variation: a) among individuals within populations, b) among populations, and c) between groups of population classified as Northern groups (Mauritania and Senegal) and Southern group (Ghana, Togo and Benin). The largest component of total genetic variation was observed within individuals nested in the population (97.7%) (Table 2.4). The variation among populations within the groups was 2%.

Table 2.4: Analysis of Molecular Variance (AMOVA) in within and between S. *aurita* populations

SOURCE OF VARIATION	VARIANCE COMPONENT	% VARIANCE
Among individuals within a sampling site	18.6	0.977
Between populations within groups	0.372	0.02
Between groups (northern and southern)	0.070	0.004

The Discriminant Analysis of Principal Component (DAPC) estimation performed to determine the posterior membership probability of each sample which was represented in a structure-like plot observed on the map (Figure 2.2) showed mixing between samples and a combination of genotypes obtained from Mauritania, Senegal, Ghana, Togo and Benin. The structure-like plot was further supported by the UPGMA dendrogram.

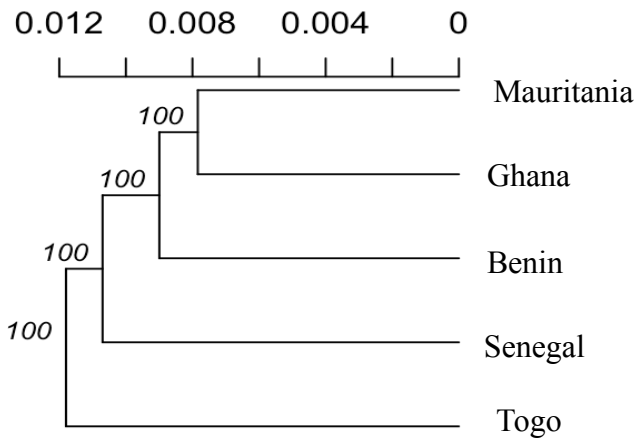

Figure 2.2: UPGMA dendrogram created using the genetic distance between the sampling sites with bootstrapped values based on 1000 replicates. Shown are all bootstrap values > =50

Sardinella maderensis: de novo assembly and SNP filtering

Samples from six populations of *S. maderensis* (Mauritania, Senegal, Togo, Guinea, Benin and Nigeria) were sequenced using double digest Rad (ddRAD) sequencing. Individuals with greater than >= 500,000 reads were used in the analysis. For *denovo* reference assembly, 61,264 contigs were obtained. Individuals (n = 122) from the six populations were initially used in mapping to the reference individual and variants calling. Nigeria samples were removed from the analysis due to extremely small number of average contigs (6,300) mapped to the reference assembly in comparison to the other populations, suggesting that the fish collected were not *S. maderensis* (supplemental table 1). After removal of Nigerian populations (20 samples), 102 individuals were retained before variants calling. We obtained raw variants calls of 985,007 SNP loci. A total of 6,767 SNPS loci passed the stringent quality threshold filtering retaining 88 individuals for downstream analysis.

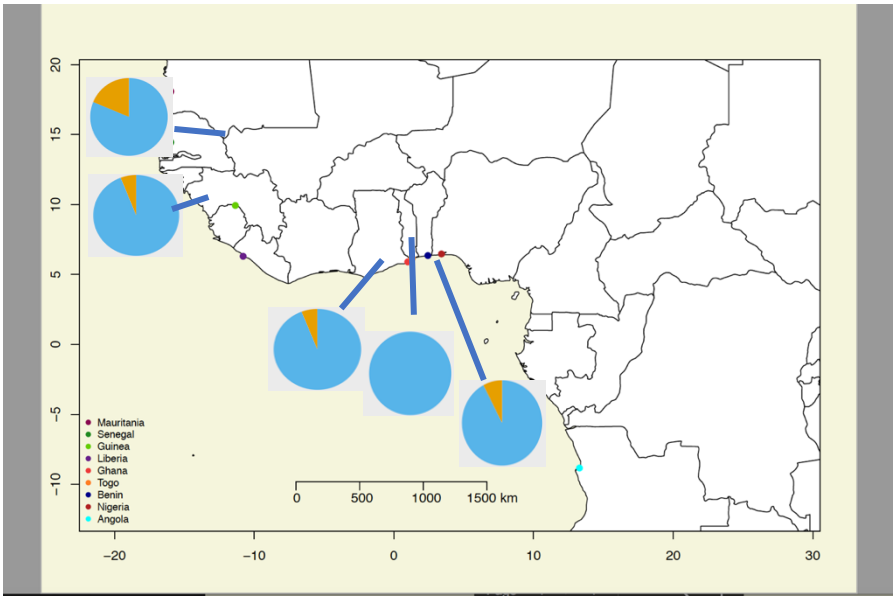

Figure 2.3: Graphical summary of structure-like plot of all individuals from the five sampling sites.

Genetic diversity in Sardinella maderensis

Genetic diversity values across the populations surveyed was estimated using the 6,767 SNP loci obtained that passed filtering. The average observed heterozygosity (Ho) values ranged from 0.321 in Senegal to 0.361 in Togo and Expected heterozygosity (He) ranged from 0.239 to 0.250 in Togo (Table 2.5). Testing of Hardy-Weinberg (HWE) equilibrium indicated no significant deviation from HWE equilibrium (not shown). Similar levels of genetic diversity were observed across all populations. The percentage of polymorphic loci ranged from 80.1% in Guinea to 97.6% in Mauritania. The F_{IS} value measured for each population ranges from -0.18 in Benin to -0.11 in Guinea, and these can be considered to be zero.

Table 2.5: Summary genetic statistics for all sampling locations. N (number of samples after SNP filtering), Sites, %polymorphic sites, Average frequency of the major allele (P), He (Expected heterozygosity) and Ho (observed heterozygosity), FIS (Wright statistics F).

Population	N	Sites	% polymorphic sites	P	Ho	He	Nucleotide diversity	FIS
Mauritania	21	6767	97.6	0.830	0.328	0.249	0.257	-0.177
Senegal	18	6767	95.0	0.835	0.321	0.239	0.247	-0.178
Guinea	10	6767	80.1	0.832	0.327	0.239	0.263	-0.135
Togo	20	6767	97.0	0.829	0.361	0.250	0.257	-0.183
Benin	19	6767	95.8	0.832	0.332	0.239	0.246	-0.189

Population structure in Sardinella maderensis

The overall F_{ST} of the full data set was 0.002, and pairwise F_{ST} values ranged from -0.005 for the Benin-Togo comparison to 0.016 for the Guinea-Togo comparison (Table 2.6). Genetic differentiation between all population comparisons was not significant (P > 0.05).

Table 2.6: Pairwise FST values calculated using 6,767 SNPs (P > 0.05)

	Mauritania	Senegal	Guinea	Togo	Benin
Mauritania					
Senegal	-0.002				
Guinea	0.000	0.006			
Togo	0.072	0.009	0.016		
Benin	0.000	0.002	-0.003	-0.005	

We conducted hierarchical Analysis of Molecular Variance (Table 2.7) for the entire data set and populations partitioned into two geographical groups; North CECAF (Mauritania, Senegal Guinea) and South CECAF (Togo, Benin) and the largest component of total genetic variation was observed within individuals nested in the population (99.8%).

The two geographic groups (North CECAF and south CECAF) accounted for 1% of the total observed variation, and the remaining variation among populations was 1%. These results also support the pairwise F_{ST} values where no genetic differentiation is observed between the population. The UPGMA dendrogram generated using the Nei genetic distance as a distance metric also shows no apparent structure in the population. As a further test of population structure in this species, the population membership probability for each (Figure 2.4) clearly shows one grouping where Mauritania, Senegal, Guinea, Benin and Togo shows as one population with strong admixture between them.

Table 2.7: Results for a standard AMOVA for the five populations of S. maderensis within two regional groups (Northern CECAF, Southern CECAF)

Source of variation	Variance component	% variance
WITHIN POPULATION	231.8	0.99
AMONG POPULATION	0.285	0.001
AMONG GROUPS	0.239	0.001

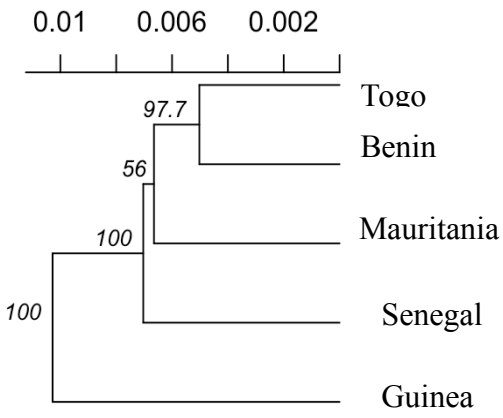

Figure 2.4: Dendrogram generated using Nei genetic distance values as distance metric. Bootstrap values at nodes greater than 50

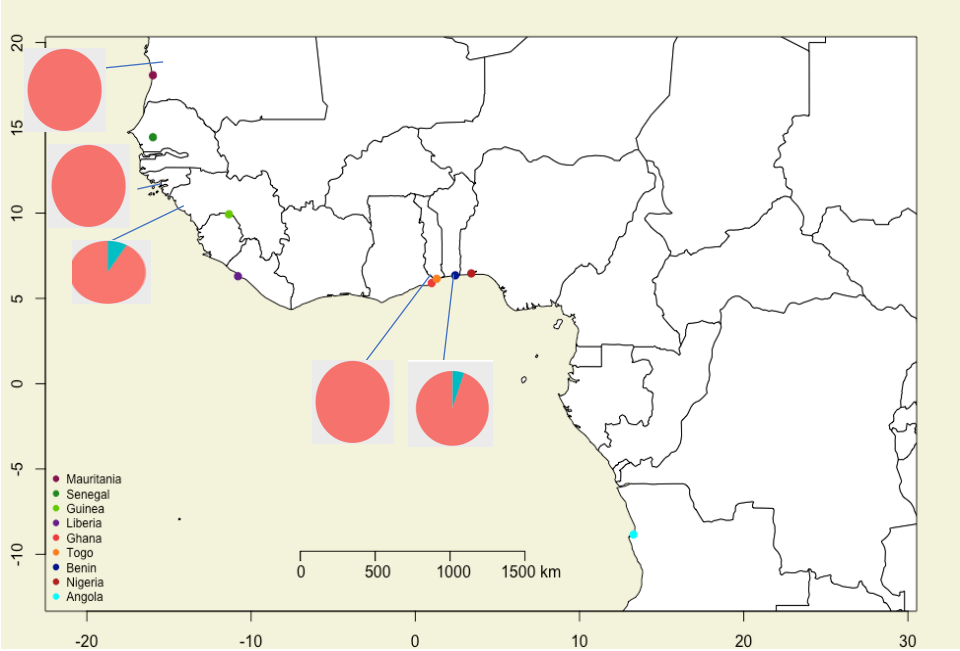

Figure 2.5: Graphical summary of membership group assignment of individuals from five populations surveyed per site scored with 6,767 SNPs.

DISCUSSION

In this study, our aim was to determine the population genetic structure of *S. aurita* and *S. maderensis* across selected localities in the coast of West Africa (Mauritania to Benin) in order to understand spatial patterns of population structure of these two sardinellas. The use of hundreds or thousands of genome-wide polymorphic markers (6,078 for *S. aurita* and 6,767 *for S. maderensis*) should allow for the detection of genetic differentiation where inferences from a single or a few marker-based inferences fail (Pukk et al. (2013). This study showed no apparent genetic differentiation and population structure in sardinella species in the coast of West Africa.

The results of this work are not unexpected for highly migratory species such as *S. aurita* and *S. maderensis,* and indicate that no barriers to gene flow exist in the region extending from Mauritania to Benin. Generally, low F_{ST} values are observed in highly migratory pelagic fishes such as sardines, sardinellas, and anchovies (Ward et al., 1994). Sukumaran et al., (2017) reported the lack of genetic subdivision between mackerel populations from India suggesting adequate gene flow and panmixia. Homogeneity among populations of Indian oil sardine (*sardinella longiceps*) populations, as detected using mitochondrial genes, has also been reported (Sukumaran, et al., 2016). Studies performed using a comparable number of markers utilized in this study (106,652 SNPs obtained using RADseq) also found a single panmictic population in the Japanese eel that lives in various environments including fresh, brackish and coastal waters from Japan (Gong et al., 2019). Dunlap, (2017) analyzed 11,836 SNPs obtained from 295 individuals of coral reef - dwelling cardinalfish, *Siphamia tubifer from Okinawa islands* from 15 locations in Japan collected within the period (2012, 2013 , 2014) and found very little genetic differentiation

between groups and no evidence of population. The low differentiation from this study can be attributed to high gene flow occurring as the result of adult migration along the coast and/or larval seaward migration and passive transport by current systems during their larval life stage, facilitating the exchange of individuals among geographic locations.

This study also confirms the results from previous studies performed with other types of markers (allozymes, microsatellites) showing no significant differentiation of *S. aurita and S. maderensis* in the CECAF and other regions. The values for overall F_{ST} of 0.001 and the pattern of the pairwise differentiation (-0.002-0.004) using RAD-SNPs in this study were comparable to those found in a previous work by (Chi, 1998) who studied genetic differentiation using allozyme markers in *S. aurita* from Congo, Ghana, Ivory Coast and Venezuela (pairwise F_{ST} between 0 - 0.0055). Our survey of 6,767 SNP loci from *S. maderensis* indicating no genetic differentiation (global F_{ST} = 0.001, $p<0.01$) is also comparable to (French et al., 1995) who previously carried out genetic differentiation of this species using allozyme markers from Senegal, Cote divoire, Ghana and Congo and found out that genetic differentiation (F_{ST} = 0.0085) exists for *S. maderensis* using allozyme markers however regarded this differentiation as to not to be biologically significant. The lack of population structure observed using RAD-SNPs is also consistent with the results of Kinsey et al. (1994) in *S. aurita* from the Gulf of Mexico to the coast of South Carolina, who studied population structure using allozyme markers and found no evidence of population structure and regarded the population to be in a state of panmixia.

S. aurita and s. maderensis are migratory fish and because of their pattern of migration (both north and south, as well as an inshore-offshore pattern of movements), there is the possibility of individual stocks to be found in the territorial waters of each of

these countries at different stages of their life cycles (Brainerd, 1991). Also the current system creates seasonal upwellings which mainly account for the distribution of the fishery resources in the region. The abundance and distribution of these species rely on variability of the coastal upwelling intensity and the associated variations in phytoplankton production (Arístegui et al., 2006). For instance, *S. aurita* is found in Senegal waters in May -June where changes in temperature causes them to migrate to Mauritania. The Ghana- Cote d'Ivoire upwelling causes the species to migrate towards Cote d'Ivoire, Togo and Benin. The absence of significant difference among locations and in allele frequencies among populations suggest gene flow occurring between these locations may neutralize population differentiation due to genetic drift or a balancing selection is maintained across these locations (Karl and Avise 1992). According to Palumbi and Baker (1994), pelagic fishes such as sardines are expected to show little panmictic subdivision because of the apparent lack of physical barriers in the marine realm, which favors a high level of egg, larva, and adult dispersion in the species, as well as greatly facilitate extensive gene flow among the populations.

One major concern in this study is the number of samples retained for downstream analysis. Sample numbers in each population (initial target was 20 per population) were limited by the number of individuals for which DNA of high enough quality for library preparation was obtained. All samples from Guinea were removed from the *S. aurita* analysis as a result of only seven individuals remaining as a representative for Guinea populations to call genotypes and determine allele frequencies, since small sample sizes can lead to an overestimation of genetic differentiation (Gong et al., 2019). Moreover, the small number of contigs mapped to the reference assembly suggested problems with DNA

quality and/or misidentification of individuals during collection. Removal of Nigeria samples from the *S. maderensis* analysis was primarily due to low coverage of individuals that mapped to the reference assembly in comparison to the rest of the localities, suggesting that the samples were misidentified as *S. maderenesis* at the landing site. This hypothesis is supported by the observation that, when included in the analysis, the Nigeria population showed high pairwise F_{ST} values (0.36-0.39, results not shown), in comparison to the other populations, values that suggest those fish belonged to a different species. The removal of these two populations and a few selected individuals from other populations during SNP filtering resulted in a reduction in the total sample size for each species (81 samples for *S. aurita*, 88 samples for *S. maderensis*). In general population genetics inferences with greater sample sizes (more individuals) representing a particular population can increase the precision in making inferences in relation to genetic differentiation by ensuring that allele frequencies detected indicate true representatives in the total population in each of the locations (Hale et al., 2012;). This study analyzed >6,000 markers in each of the species using between 15 to 20 individuals obtained from each of the locations. Although this number of samples as a representative of each location may be small, population differentiation values such as F_{ST} can be accurately calculated for small sample sizes, given that the number of biallelic markers examined is high (n > 1000) (Willing et al., 2012).

As effective fishery management requires to correctly quantify the connectivity patterns among stocks, genetic studies are of much importance as they provide strength to indicate differentiation in species. The genetic analysis of *S. aurita* and *S. maderensis* has shown no subdivision in populations obtained from Mauritania to Benin and did not

support the division of Guinea as a Northern stock, Mauritania, Senegal classified as to be in Northern CECAF and Ghana, Togo, Benin as Western stock. Based on these results, fish in this region should be considered as belonging to a single panmictic population. Samples were not obtained for central stock (Nigeria and Cameroon) and southern stock (Gabon, the Democratic Republic of the Congo, the Congo and Angola), so this study should be expanded to include these regions. These results will therefore enhance the decision making towards the management of these species in the Eastern Central Atlantic region for sustainability and form a baseline for future analysis of this populations. Moreover, this work has provided sequence information and assemblies for two sardinella species that can be used in the future to further develop other genotyping platforms, through mining of microsatellite markers or panels of informative SNPs.

Population genetics inferences have the potential to increase accuracy with greater sample sizes which can also increase the precision of allele frequency estimates as a representative of the population within that locality. Depending on the number of markers used in future studies, I will therefore recommend to incorporate larger sample sizes to ensure that most informative alleles can be sampled at frequencies that reflect those in the total population (Hale et al., 2012). Also, samples collected from different years can asist in estimation of changes in allele frequencies with respect to different temporal scale (Dunlap, 2017). I will also recommend sampling from different time frames (year) to further ascertain the pattern of differentiation spatially and temporally because the dynamics of temporal genetic structure may even be more informative than spatial dynamics in marine populations (Hedgecock et al., 2007).

CONCLUSION

This study detected low levels of genetic differentiation and lack of population structure in the *S. aurita* and *S. maderensis* from the geographic locations surveyed, suggesting the existence of one population from Mauritania, Senegal, Ghana, Togo and Benin as compared to the existence of subdivision proposed by the CECAF working group for either of the species.

3 APPENDIX

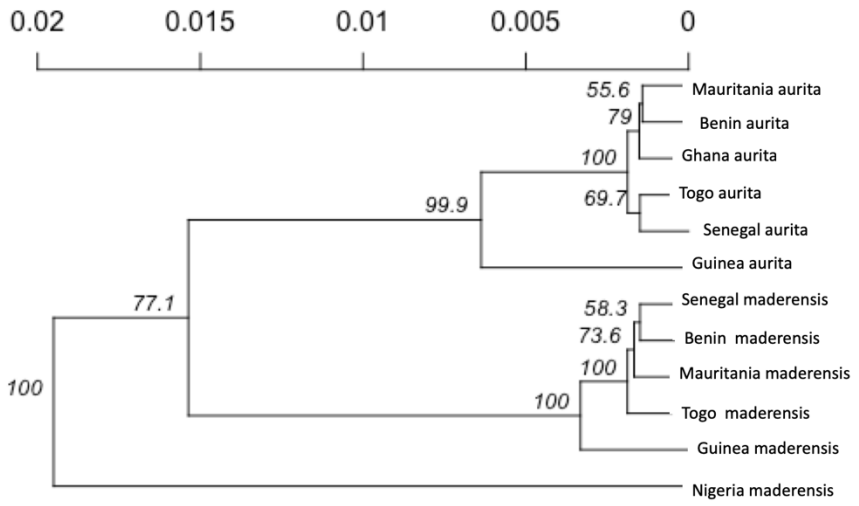

Figure 3.1: Dendrogram plot generated to differentiate between S. *aurita* and S. *maderensis*

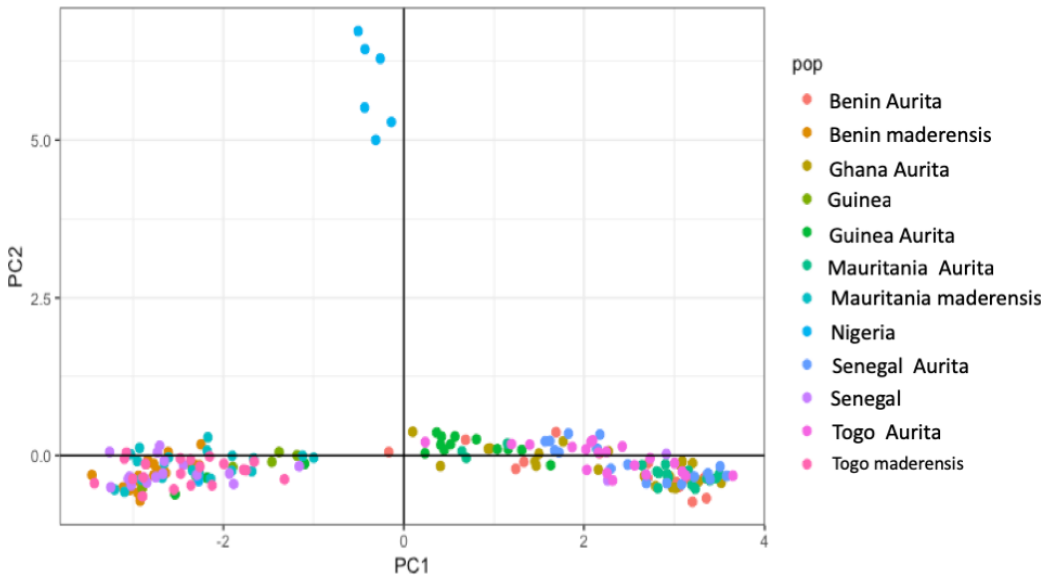

Figure 3.2: PCA plot generated to differentiate between S. *aurita* and S. *maderensis*.

Table 3.1: Results of Mapping to reference assembly for both s. aurita and s. maderensis

	S. aurita			***S. maderensis***	
Population	Average contig mapped	Average coverage	Population	Average contig mapped	Average coverage
Mauritania	42623	6.5	Mauritania	17788	4.7
Senegal	59678	6.4	Senegal	24145	6.8
Guinea	17283	4.6	Guinea	7231	4.5
Ghana	36637	7.2	Benin	28674	6.5
Benin	37741	6.3	Togo	17662	6.8
Togo	36825	6.9	Nigeria	6538	3.4

Table 3.2: SNP filtering procedure

Filtering steps	***S. aurita***	***S. maderensis***
Initial Variants calls	1256073 sites in the raw variants and 95 individuals	[1]985007 sites in the raw variants and 102 individuals
Variants that have been successfully genotyped in 50% of individuals with a minimum quality score of 20, and a minor allele count of 3	kept 449662 out of a possible 1256073 Sites,	[2]kept 157860 out of a possible 985007 sites
loci with a minimum depth less than 5 and minor allele frequency less than 1%	kept 353136 out of a possible 449662 Sites	[3]kept 145131 out of 157860 Sites
individuals with 60% missing data filtered using the script (filter_missing_ind.sh)	kept 81 individuals out of 94 and 353136 loci.	[4]kept 88 out of 102 Individuals and 145131 loci
The script pop_missing_filter.sh to filter out loci by population with 90% missing data.	kept 353136 out of a possible 353136 Sites	[5]kept 69019 out of a possible 145131 Sites
Filtering using dDocent_filters script	174199 sites were kept out of 353136 sites	
INDELS removed using VCFtools to produce a VCF file of only SNP calls	14199 out of a possible 174199 Sites.	[7]kept 39054 out of a possible 42218 Sites
bash script filter filter_hwe_by_pop.pl was deviated from HWE in each population		[8]Kept 38461 out of a possible 39054
bash script rad_haplotyper.pl filter	kept 6155 loci out of 14199	
Loci with minor allele frequency threshold maf < 0.05	kept 6155 out of a possible 6155 Sites	[6]kept 42218 out of a possible 69019 Sites.
genotype call rate of 75% across all individuals		[9]kept 7106 out of a possible 38461 Sites
SNPs filtered to include loci with two alleles	kept 6155 out of a possible 6155 Sites	[10]6767 sites out of a possible 7106 Sites were retained
Bayescan outlier detection	Kept 6078 neutral loci out of possible 6081	[11]Kept 6,767 neutral loci out of possible 6767

Table 3.3: HiSeq Read Set Sequencing Metadata

Sample	country	species	Run Type	Library Type	Read #	Average Quality
MA_D1	Mauritania	S. aurita	PAIRED_END	ddRAD	3972953	40
MA_D2	Mauritania	S. aurita	PAIRED_END	ddRAD	2863133	40
MA_D3	Mauritania	S. aurita	PAIRED_END	ddRAD	3836010	40
MA_D7	Mauritania	S. aurita	PAIRED_END	ddRAD	2997783	40
MA_D8	Mauritania	S. aurita	PAIRED_END	ddRAD	3762103	40
MA_D9	Mauritania	S. aurita	PAIRED_END	ddRAD	3029751	40
MA_G1	Mauritania	S. aurita	PAIRED_END	ddRAD	3114856	40
MA_G2	Mauritania	S. aurita	PAIRED_END	ddRAD	3448832	40
MA_G3	Mauritania	S. aurita	PAIRED_END	ddRAD	2990393	40
MA_G4	Mauritania	S. aurita	PAIRED_END	ddRAD	3363135	40
MA_G5	Mauritania	S. aurita	PAIRED_END	ddRAD	2326233	40
MA_G6	Mauritania	S. aurita	PAIRED_END	ddRAD	2888975	40
MA_G8	Mauritania	S. aurita	PAIRED_END	ddRAD	4453339	40
MA_G10	Mauritania	S. aurita	PAIRED_END	ddRAD	3317070	40
MA_1A	Mauritania	S. aurita	PAIRED_END	ddRAD	2388069	40
MA_2A	Mauritania	S. aurita	PAIRED_END	ddRAD	498866	40
MA_7A	Mauritania	S. aurita	PAIRED_END	ddRAD	440287	40
MA_8A	Mauritania	S. aurita	PAIRED_END	ddRAD	426360	40
MA_9A	Mauritania	S. aurita	PAIRED_END	ddRAD	1423744	40
SA_C1	Senegal	S. aurita	PAIRED_END	ddRAD	2909942	40
SA_C2	Senegal	S. aurita	PAIRED_END	ddRAD	2322853	40
SA_C3	Senegal	S. aurita	PAIRED_END	ddRAD	3231356	40
SA_C4	Senegal	S. aurita	PAIRED_END	ddRAD	3268838	40
SA_C5	Senegal	S. aurita	PAIRED_END	ddRAD	2507019	40
SA_C6	Senegal	S. aurita	PAIRED_END	ddRAD	6007766	40
SA_C7	Senegal	S. aurita	PAIRED_END	ddRAD	2334828	40
SA_C8	Senegal	S. aurita	PAIRED_END	ddRAD	2426209	36
SA_C9	Senegal	S. aurita	PAIRED_END	ddRAD	2855746	40
SA_C10	Senegal	S. aurita	PAIRED_END	ddRAD	3235321	40
SA_1C	Senegal	S. aurita	PAIRED_END	ddRAD	1829026	40
SA_2C	Senegal	S. aurita	PAIRED_END	ddRAD	2884170	40
SA_3C	Senegal	S. aurita	PAIRED_END	ddRAD	2374320	40
SA_5Cr	Senegal	S. aurita	PAIRED_END	ddRAD	67387	40
SA_7Cr	Senegal	S. aurita	PAIRED_END	ddRAD	1490195	40
SA_8Cr	Senegal	S. aurita	PAIRED_END	ddRAD	3220885	40

SA_38	Senegal	S. aurita	PAIRED_END	ddRAD	1807230	40
SA_4CN	Senegal	S. aurita	PAIRED_END	ddRAD	3490025	40
GUI_11	Guinea	S. aurita	PAIRED_END	ddRAD	1671444	40
GUI_13	Guinea	S. aurita	PAIRED_END	ddRAD	2736320	40
GUI_14	Guinea	S. aurita	PAIRED_END	ddRAD	3281360	40
GUI_15	Guinea	S. aurita	PAIRED_END	ddRAD	3738207	40
GUI_16	Guinea	S. aurita	PAIRED_END	ddRAD	2614403	40
GUI_19	Guinea	S. aurita	PAIRED_END	ddRAD	2539545	40
GUI_1G	Guinea	S. aurita	PAIRED_END	ddRAD	2416160	40
GUI_1H	Guinea	S. aurita	PAIRED_END	ddRAD	1950201	40
GUI_1	Guinea	S. aurita	PAIRED_END	ddRAD	2096833	40
GUI_21	Guinea	S. aurita	PAIRED_END	ddRAD	3559740	40
GUI_22	Guinea	S. aurita	PAIRED_END	ddRAD	2235137	40
GUI_23	Guinea	S. aurita	PAIRED_END	ddRAD	2786868	40
GUI_25	Guinea	S. aurita	PAIRED_END	ddRAD	4059001	40
GUI_2	Guinea	S. aurita	PAIRED_END	ddRAD	3573740	40
GUI_31	Guinea	S. aurita	PAIRED_END	ddRAD	3453429	40
GUI_37	Guinea	S. aurita	PAIRED_END	ddRAD	778994	40
GUI_43	Guinea	S. aurita	PAIRED_END	ddRAD	1068216	40
GUI_11	Guinea	S. aurita	PAIRED_END	ddRAD	1644856	40
GUI_4G	Guinea	S. aurita	PAIRED_END	ddRAD	654806	40
GUI_5H	Guinea	S. aurita	PAIRED_END	ddRAD	1233015	40
GA	Ghana	S. aurita	PAIRED_END	ddRAD	4682704	40
GA_A4	Ghana	S. aurita	PAIRED_END	ddRAD	3296860	40
GA_A2	Ghana	S. aurita	PAIRED_END	ddRAD	3274985	40
GA_A3	Ghana	S. aurita	PAIRED_END	ddRAD	176859	40
GA_A5	Ghana	S. aurita	PAIRED_END	ddRAD	2975779	40
GA_A6	Ghana	S. aurita	PAIRED_END	ddRAD	5448638	40
GA_A7	Ghana	S. aurita	PAIRED_END	ddRAD	3521271	40
GA_A8	Ghana	S. aurita	PAIRED_END	ddRAD	2593197	40
GA_A9	Ghana	S. aurita	PAIRED_END	ddRAD	2336965	40
GA_A10	Ghana	S. aurita	PAIRED_END	ddRAD	3822138	40
GA_F6	Ghana	S. aurita	PAIRED_END	ddRAD	1753505	40
GA_F7	Ghana	S. aurita	PAIRED_END	ddRAD	7250197	40
GA_F8	Ghana	S. aurita	PAIRED_END	ddRAD	1904330	40
GA_F9	Ghana	S. aurita	PAIRED_END	ddRAD	2578087	40
GA_F10	Ghana	S. aurita	PAIRED_END	ddRAD	1791180	40
GA_1B	Ghana	S. aurita	PAIRED_END	ddRAD	777363	40
GA_2B	Ghana	S. aurita	PAIRED_END	ddRAD	1051070	40

GA_3B	Ghana	S. aurita	PAIRED_END	ddRAD	760786	40
GA_5B	Ghana	S. aurita	PAIRED_END	ddRAD	1085310	40
GA_gha	Ghana	S. aurita	PAIRED_END	ddRAD	1559745	40
TA_E1	Togo	S. aurita	PAIRED_END	ddRAD	1903707	40
TA_E2	Togo	S. aurita	PAIRED_END	ddRAD	2920657	40
TA_E3	Togo	S. aurita	PAIRED_END	ddRAD	6364394	40
TA_E4	Togo	S. aurita	PAIRED_END	ddRAD	2399133	40
TA_E5	Togo	S. aurita	PAIRED_END	ddRAD	2825722	40
TA_E6	Togo	S. aurita	PAIRED_END	ddRAD	2330258	40
TA_E7	Togo	S. aurita	PAIRED_END	ddRAD	2163294	40
TA_E8	Togo	S. aurita	PAIRED_END	ddRAD	4392504	40
TA_E9	Togo	S. aurita	PAIRED_END	ddRAD	3699474	40
TA_E10	Togo	S. aurita	PAIRED_END	ddRAD	1982836	40
TA_41	Togo	S. aurita	PAIRED_END	ddRAD	3335789	40
TA_44	Togo	S. aurita	PAIRED_END	ddRAD	3130346	40
TA_45	Togo	S. aurita	PAIRED_END	ddRAD	3762442	40
TA_47	Togo	S. aurita	PAIRED_END	ddRAD	3393171	40
TA_1Dre	Togo	S. aurita	PAIRED_END	ddRAD	669464	40
TA_2Dre	Togo	S. aurita	PAIRED_END	ddRAD	2721236	40
TA_3Dre	Togo	S. aurita	PAIRED_END	ddRAD	652961	40
TA_4D	Togo	S. aurita	PAIRED_END	ddRAD	2033845	40
TA_5D	Togo	S. aurita	PAIRED_END	ddRAD	1966840	40
TA_6Dr	Togo	S. aurita	PAIRED_END	ddRAD	3079161	40
BA_1E	Benin	S. aurita	PAIRED_END	ddRAD	648405	40
BA_2E	Benin	S. aurita	PAIRED_END	ddRAD	3250561	40
BA_3E	Benin	S. aurita	PAIRED_END	ddRAD	1175269	40
BA_4E	Benin	S. aurita	PAIRED_END	ddRAD	1138702	40
BA_5E	Benin	S. aurita	PAIRED_END	ddRAD	426831	36
BA_6E	Benin	S. aurita	PAIRED_END	ddRAD	1160370	40
BA_7E	Benin	S. aurita	PAIRED_END	ddRAD	1864077	40
BA_B1	Benin	S. aurita	PAIRED_END	ddRAD	421762	37
BA_B10	Benin	S. aurita	PAIRED_END	ddRAD	3516024	40
BA_B2	Benin	S. aurita	PAIRED_END	ddRAD	3049783	40
BA_B3	Benin	S. aurita	PAIRED_END	ddRAD	2699346	40
BA_B4	Benin	S. aurita	PAIRED_END	ddRAD	2277126	40
BA_B5	Benin	S. aurita	PAIRED_END	ddRAD	2994343	36
BA_B6	Benin	S. aurita	PAIRED_END	ddRAD	2529279	40
BA_B7	Benin	S. aurita	PAIRED_END	ddRAD	3737002	40
BA_B8	Benin	S. aurita	PAIRED_END	ddRAD	6323212	40

BA_B9	Benin	S. aurita	PAIRED_END	ddRAD	6343378	40
MM10A	Mauritania	maderensis	PAIRED_END	ddRAD	2429179	40
MM11A	Mauritania	maderensis	PAIRED_END	ddRAD	604046	40
MM11B	Mauritania	maderensis	PAIRED_END	ddRAD	2221548	40
MM12A	Mauritania	maderensis	PAIRED_END	ddRAD	2773836	40
MM_1A	Mauritania	maderensis	PAIRED_END	ddRAD	1002115	36
MM_2A	Mauritania	maderensis	PAIRED_END	ddRAD	707037	36
MM_3A	Mauritania	maderensis	PAIRED_END	ddRAD	1770842	36
MM_4A	Mauritania	maderensis	PAIRED_END	ddRAD	1225941	36
MM_4B	Mauritania	maderensis	PAIRED_END	ddRAD	1988491	36
MM_5A	Mauritania	maderensis	PAIRED_END	ddRAD	1001721	36
MM_6A	Mauritania	maderensis	PAIRED_END	ddRAD	1020230	36
MM_6B	Mauritania	maderensis	PAIRED_END	ddRAD	964148	36
MM_7A	Mauritania	maderensis	PAIRED_END	ddRAD	1096085	36
MM_7B	Mauritania	maderensis	PAIRED_END	ddRAD	3315913	36
MM_8A	Mauritania	maderensis	PAIRED_END	ddRAD	1383283	36
MM_9A	Mauritania	maderensis	PAIRED_END	ddRAD	3834032	36
MM_9B	Mauritania	maderensis	PAIRED_END	ddRAD	2553159	40
MMD10	Mauritania	maderensis	PAIRED_END	ddRAD	2221548	40
MMD4	Mauritania	maderensis	PAIRED_END	ddRAD	3620659	40
SM__1B	Senegal	maderensis	PAIRED_END	ddRAD	2084735	40
SM_1C	Senegal	maderensis	PAIRED_END	ddRAD	1317125	40
SM__2B	Senegal	maderensis	PAIRED_END	ddRAD	468068	40
SM__2C	Senegal	maderensis	PAIRED_END	ddRAD	2394435	40
SM__3B	Senegal	maderensis	PAIRED_END	ddRAD	1340907	40
SM__3C	Senegal	maderensis	PAIRED_END	ddRAD	1246669	40
SM_4C	Senegal	maderensis	PAIRED_END	ddRAD	2197183	40
SM__4C	Senegal	maderensis	PAIRED_END	ddRAD	639825	40
SM__5C	Senegal	maderensis	PAIRED_END	ddRAD	3494326	40
SM_H10	Senegal	maderensis	PAIRED_END	ddRAD	2025534	40
SM_H1	Senegal	maderensis	PAIRED_END	ddRAD	1209463	36
SM_H2	Senegal	maderensis	PAIRED_END	ddRAD	2934705	36
SM_H3	Senegal	maderensis	PAIRED_END	ddRAD	2786505	36
SM_H4	Senegal	maderensis	PAIRED_END	ddRAD	3530385	40
SM_H5	Senegal	maderensis	PAIRED_END	ddRAD	4799496	40
SM_H6	Senegal	maderensis	PAIRED_END	ddRAD	3258402	40
SM_H7	Senegal	maderensis	PAIRED_END	ddRAD	3135016	40
SM_H8	Senegal	maderensis	PAIRED_END	ddRAD	3045871	40
SM_H9	Senegal	maderensis	PAIRED_END	ddRAD	3357428	40

GM_10E	Senegal	maderensis	PAIRED_END	ddRAD	2814198	40
GM_23	Guinea	maderensis	PAIRED_END	ddRAD	2162394	40
GM_25	Guinea	maderensis	PAIRED_END	ddRAD	2579701	40
GM_26	Guinea	maderensis	PAIRED_END	ddRAD	2538471	40
GM_27	Guinea	maderensis	PAIRED_END	ddRAD	5642128	40
GM_28	Guinea	maderensis	PAIRED_END	ddRAD	5698079	36
GM_29	Guinea	maderensis	PAIRED_END	ddRAD	3328832	36
GM_30	Guinea	maderensis	PAIRED_END	ddRAD	3586765	36
GM_33	Guinea	maderensis	PAIRED_END	ddRAD	4175486	36
GM_34	Guinea	maderensis	PAIRED_END	ddRAD	3136132	36
GM_35	Guinea	maderensis	PAIRED_END	ddRAD	3974891	40
GM_36	Guinea	maderensis	PAIRED_END	ddRAD	3757137	40
GM_37	Guinea	maderensis	PAIRED_END	ddRAD	3116255	40
GM_38	Guinea	maderensis	PAIRED_END	ddRAD	4791637	40
GM_39	Guinea	maderensis	PAIRED_END	ddRAD	2728215	40
GM_40	Guinea	maderensis	PAIRED_END	ddRAD	3664194	40
GM_6E	Guinea	maderensis	PAIRED_END	ddRAD	4389357	40
GM_6F	Guinea	maderensis	PAIRED_END	ddRAD	3930490	36
GM_7E	Guinea	maderensis	PAIRED_END	ddRAD	3393483	36
GM_8E	Guinea	maderensis	PAIRED_END	ddRAD	3026438	40
N1	Nigeria	maderensis	PAIRED_END	ddRAD	3798103	36
N2	Nigeria	maderensis	PAIRED_END	ddRAD	2818093	36
N3	Nigeria	maderensis	PAIRED_END	ddRAD	1847396	40
N4	Nigeria	maderensis	PAIRED_END	ddRAD	1882335	40
N5	Nigeria	maderensis	PAIRED_END	ddRAD	1458661	36
N6	Nigeria	maderensis	PAIRED_END	ddRAD	2077416	40
N7	Nigeria	maderensis	PAIRED_END	ddRAD	4632541	40
N8	Nigeria	maderensis	PAIRED_END	ddRAD	3651073	40
N9	Nigeria	maderensis	PAIRED_END	ddRAD	1764589	40
N10	Nigeria	maderensis	PAIRED_END	ddRAD	863906	40
N11	Nigeria	maderensis	PAIRED_END	ddRAD	532017	40
N12	Nigeria	maderensis	PAIRED_END	ddRAD	2093608	36
N13	Nigeria	maderensis	PAIRED_END	ddRAD	4897107	36
N14	Nigeria	maderensis	PAIRED_END	ddRAD	2095123	36
N15	Nigeria	maderensis	PAIRED_END	ddRAD	432852	40
N16	Nigeria	maderensis	PAIRED_END	ddRAD	2996849	36
N17	Nigeria	maderensis	PAIRED_END	ddRAD	1436918	36
N18	Nigeria	maderensis	PAIRED_END	ddRAD	562976	40
N19	Nigeria	maderensis	PAIRED_END	ddRAD	1859757	36

N20	Nigeria	maderensis	PAIRED_END	ddRAD	4351480	40
TM_10B	Togo	maderensis	PAIRED_END	ddRAD	1490224	40
TM_11A	Togo	maderensis	PAIRED_END	ddRAD	2093608	40
TM_11B	Togo	maderensis	PAIRED_END	ddRAD	1318183	40
TM_12A	Togo	maderensis	PAIRED_END	ddRAD	3250561	40
TM_12B	Togo	maderensis	PAIRED_END	ddRAD	1546825	40
TM_26	Togo	maderensis	PAIRED_END	ddRAD	745465	40
TM_37	Togo	maderensis	PAIRED_END	ddRAD	498351	40
TM_38	Togo	maderensis	PAIRED_END	ddRAD	2610747	40
TM_39	Togo	maderensis	PAIRED_END	ddRAD	1490224	40
TM_3A	Togo	maderensis	PAIRED_END	ddRAD	3150474	40
TM_45	Togo	maderensis	PAIRED_END	ddRAD	6004595	40
TM_4A	Togo	maderensis	PAIRED_END	ddRAD	485764	40
TM_4B	Togo	maderensis	PAIRED_END	ddRAD	1165329	40
TM_5A	Togo	maderensis	PAIRED_END	ddRAD	2399695	40
TM_5B	Togo	maderensis	PAIRED_END	ddRAD	769184	36
TM_6A	Togo	maderensis	PAIRED_END	ddRAD	4202676	36
TM_6B	Togo	maderensis	PAIRED_END	ddRAD	3721938	40
TM_8A	Togo	maderensis	PAIRED_END	ddRAD	3230095	36
TM_8B	Togo	maderensis	PAIRED_END	ddRAD	4299789	40
TM_8B	Togo	maderensis	PAIRED_END	ddRAD	3032301	40
BM_10E	Benin	maderensis	PAIRED_END	ddRAD	3345083	40
BM_11E	Benin	maderensis	PAIRED_END	ddRAD	2280475	40
BM_12E	Benin	maderensis	PAIRED_END	ddRAD	1977142	36
BM_1F	Benin	maderensis	PAIRED_END	ddRAD	3459880	36
BM_2F	Benin	maderensis	PAIRED_END	ddRAD	2637252	36
BM_3F	Benin	maderensis	PAIRED_END	ddRAD	1915437	36
BM_4F	Benin	maderensis	PAIRED_END	ddRAD	1858038	36
BM_5F	Benin	maderensis	PAIRED_END	ddRAD	5330703	36
BM_7F	Benin	maderensis	PAIRED_END	ddRAD	1415566	40
BM_A11	Benin	maderensis	PAIRED_END	ddRAD	1508411	40
BM_A12	Benin	maderensis	PAIRED_END	ddRAD	5986337	40
BM_B11	Benin	maderensis	PAIRED_END	ddRAD	3322288	36
BM_B12	Benin	maderensis	PAIRED_END	ddRAD	5785770	36
BM_C11	Benin	maderensis	PAIRED_END	ddRAD	3084180	36
BM_D11	Benin	maderensis	PAIRED_END	ddRAD	2360582	36
BM_E11	Benin	maderensis	PAIRED_END	ddRAD	3289491	36
BM_F11	Benin	maderensis	PAIRED_END	ddRAD	2817261	36
BM_G11	Benin	maderensis	PAIRED_END	ddRAD	3926321	36

BM_H11	Benin	maderensis	PAIRED_END	ddRAD	3143584	36

4 BIBLIOGRAPHY

Amponsah, S. K. K., & Ofori-danson, P. K., 2019. Estimates of Population parameters for *Sardinella maderensis* (Lowe , 1838) in the coastal waters of Ghana, (February). https://doi.org/10.15580/GJAS.2019.1.011719017.

Atarhouch T, Rüber L, Gonzalez EG, Albert EM, Rami M,, 2006. Signature of an early genetic bottleneck in a population of Moroccan sardines (*Sardina pilchardus*). Mol Phylogenet Evol 39: 373-383. doi: 10.1016/j.ympev.2005.08.003. PubMed: 16216537.

Baird, N. A., P. D., Etter, T. S., Atwood, M. C., Currey, A. L., Shiver, Z. A., Lewis, E. U., Selker, et al., 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. PLoS ONE 3:e3376.

Benestan, L., Gosselin, T., Perrier, C., 2015. RAD-genotyping reveals fine-scale genetic structuring and provides powerful population assignment in a widely distributed marine species; the American lobster (Homarus americanus). Mol Ecol.

Bernatchez, L., and C. C. Wilson. 1998. Comparative phylogeography of Nearctic and Palearctic fishes. Mol. Ecol. 7:431–451.

Belvèze, H. 1991. Influence des facteurs hydroclimatiques sur la pêcherie marocaine de petits pélagiques côtiers. In Pêcheries Ouest Africaines: Variabilité, Instabilité et Changement. Cury, P. and Roy, C. (Editors): ORSTOM, Paris: 209-233.

Brainerd, T.R. (1991). The Sardinella Fishery off the Coast of West Africa; The Case of A Common Property Resource. Paper presented at the Second Annual Conference of the International Association for the Study of Common Property (IASCP), September 26-29, 1991, University of Manitoba, Winnipeg, Canada. 34 pp.

Bradbury, I. R., S. Hubert, B. Higgins, S. Bowman, T. Borza, I. G. Paterson, P. V. R. Snelgrove et al. (2013) Genomic islands of divergence and their consequences for the

resolution of spatial structure in an exploited marine fish. Evolutionary Applications 6:450–461.

Boely, T. and Freon, P., (1979). Les ressources en poissons pelagiques des Cotes ouest-africaines de la Mauritanie au Congo. FAO, Rome.

Boëly, T., Chabanne, J., Fréon, P. & Stéquert, B. (1982). Cycle sexuel et migrations de *S. aurita* sur le plateau continental ouest-africain des îles Bissagos à la Mauritanie. *Rapport proces-verbaux des Réunion – Conseil international pour l'Exploration de la Mer***180**, 350– 355.

Bourret, V., Kent, M. P., Primmer, C.R., Vasem€agi, A., Karlsson, S., Hindar, K., McGinnity, P., 2013. SNP-array reveals genome-wide patterns of geographical and potential adaptive divergence across the natural range of Atlantic salmon (*Salmo salar*). Molecular Ecology 22:532–551.

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Genome analysis Trimmomatic : a flexible trimmer for Illumina sequence data, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Brennan, R.S., Hwang, R., Tse, M., Fangue, N.A., Whitehead, A. Local adaptation to osmotic environment in killifish, Fundulus heteroclitus, is supported by divergence in swimming performance but not by differences in excess post-exercise oxygen consumption or aerobic scope Comp. Biochem. Physiol. A. Mol. Integr. Physiol., 196 (2016), pp. 11-19.

Bureau, I., & Resources, F. O. R. A., 1999. Round sardinella (*Sardinella aurita*, Valenciennes, 1847) & Flat sardinella (*Sardinella maderensis*, Lowe, 1839), (Fig 1), 1 3.

Candy,J.R. ,Campbell, N.R.,Grinnell,M.H., Beacham, T.D., Larson, W.A., Narum, S.R. Population differentiation determined from putative neutral and divergent adaptive genetic markers in Eulachon (*Thaleichthys pacificus*, Osmeridae), an anadromous Pacific smelt Mol. Ecol. Res., 15 (6) (2015), pp. 1421-1434, 10.1111/1755-0998.12400.

Carvalho, G. R. & Hauser, L. 1995. Genetic impacts of fish introductions: a perspective on African lakes. In The Impact of Species Changes in African Lakes (Pitcher, T. J. & Hart, P. J. B., eds), pp. 457–494. London: Chapman & Hall.

Catchen, J. M., Amores, A., Hohenlohe, P. A., Cresko, W., & Postlethwait, J. H. (2011). Stacks: Building and genotyping loci de novo from short-read sequences. G3: Genes, Genomes, Genetics, 1, 171–182.

Chi, D. (1998). Low genetic variability in a widely distributed and abundant clupeid species , *Sardinella aurita* . New empirical results, 44, 861–878.

Chong Z, Ruan J, Wu C-I. 2012. Rainbow: an integrated tool for efficient clustering and assembling RAD-seq reads. *Bioinformatics* 28.21:2732-2737

Cury, P. & Roy, C. (1989). Optimal environmental window and pelagic fish recruitment success in upwelling areas. Canadian Journal of Fisheries and Aquatic Sciences 46, 670–680.

Davey, J. L., & Blaxter, M. W. 2010. RADseq: Next-generation population genetics. Briefings in Functional Genomics, 9(5–6), 416–423.

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, McVean G, Durbin R. The variant call format and VCFtools. Bioinformatics. 2011;27:2156–2158. doi: 10.1093/bioinformatics/btr330.

DeHaan, P. W., Libants, S. V., Elliott, R. F., and Scribner, K. T., 2006. Genetic population Structure of remnant lake sturgeon populations in the upper Great Lakes basin. Transactions of the American Fisheries Society 135:1478–1492.

Dichmont, C.M., Ovenden, J.R., Berry, O., Welch, D., Buckworth, R.C., (2012) Scoping current and future genetic tools, their limitations and their applications for wild fisheries management CSIRO, Brisbane.

Dunlap, P. V. (2017). Genomic analysis of a cardinalfish with larval homing potential reveals genetic admixture in the Okinawa Islands, (April), 3870–3882. https://doi.org/10.1111/mec.14169.

Durand, J. D., Guinand, B., Dodson, J. J., & Lecomte, F. (2013). Pelagic Life and Depth: Coastal Physical Features in West Africa Shape the Genetic Structure of the Bonga Shad, *Ethmalosa fimbriata. PLoS ONE*, 8(10), 1–14.

Excoffier, L., Smouse, P. E., & Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics* 131, 479–491.

Failler, P., 2014 Journal of Fisheries & Livestock Production, 2(2).

FAO 2009b. The state of world fisheries and aquaculture 2008. Rome: FAO.

French, C., Centre, N., Universit, F. B., View, D. N. A., Elephant, P., Genetics, C., & Bonhomme, F., 1995. The search for genetic differentiation of two sardine species (*Sardinella aurita* and *S. maderensis*).

Funk, W.C., McKay, J.K., Hohenlohe, P.A., Allendorf, F.W., (2012) Harnessing genomics for delineating conservation units. Trends in Ecology & Evolution 27, 489-496.

Freon, P. 1988. Reponses et adaptations des stocks de Clupeid& d'Afrique de l'ouest B la variabilit6 du milieu et de l'exploitatioa Analyse et reflexion B partir de l'exemple du Senkgal. Etudes et ThBses. ORSTOM.

Fu L, Niu B, Zhu Z, Wu S, Li W. CD-HIT: accelerated for clustering the next-generation Sequencing data. Bioinformatics. 2012;28:3150–3152.doi: 10.1093/bioinformatics/bts565.

Foll, M. & Gaggiotti, O. A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics* 180, 977–993 (2008).

Garrison E, Marth G. Haplotype-based variant detection from short-read sequencing. 20121207.3907.

Gong, X., Davenport, E. R., Wang, D., & Clark, A. G. (2019). Lack of spatial and temporal genetic structure of Japanese eel (Anguilla japonica) populations. *Conservation Genetics*, *0*(0), 0. https://doi.org/10.1007/s10592-019-01146-8.

Hale, M. L., Burg, T. M., & Steeves, T. E. 2012. Sampling for Microsatellite-Based Population Genetic Studies : 25 to 30 Individuals per Population Is Enough to Accurately Estimate Allele Frequencies, *7*(9).

Hess JE, Campbell NR, Close DA, Docker MF, Narum SR, 2013. Population genomics of Pacific lamprey: adaptive variation in a highly dispersive species. Molecular Ecology, 22, 2898–2916.

Hedgecock, D., Barber, P. H., & Edmands, S. 2007. Genetic approaches to measuring connectivity. Oceanography, 20, 70–79.

Hedrick, P. 2005. 'Genetic restoration': a more comprehensive perspective than 'genetic rescue'. Trends Ecol. Evol. 20: 109.

Hohenlohe PA, Bassham S, Etter PD, Stiffler N, Johnson EA, Cresko WA. Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. PLoS Genetics. 2010;6:e431. doi: 10.1371/journal.pgen.1000862.

Jombart, T., 2008. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24:1403–1405.

Keenan,K., McGinnity,P., Cross, T.F., Crozier, W.W., Prodöhl, P.A., 2013. diveRsity: an R package for the estimation and exploration of population genetics parameters and their associated errors. Methods in Ecology and Evolution, 4(8): 782-788.

Kinsey, S. T., Orsoy, T., Bert, T. M. & Mahmoudi, B., 1994. Population structure of the Spanish sardine *Sardinella aurita*: natural morphological variation in a genetically

homogeneous population. Marine Biology 118, 309–317.

Kritzer, J. P. & Sale, P. F. In Marine metapopulations (eds. JP, K. & Sale, P. F.) 31–67 Elsevier Academic Press, Burlington, MA, 2006.

Lamichhaney, S., Martinez Barrio, A., Rafati, N., 2012. Population-scale sequencing reveals genetic differentiation due to local adaptation in Atlantic herring. Proc Natl Acad Sci U S A 109, 19345-19350.

Lazar, N., Yankson K., Blay, J., Ofori-Danson, P., Markwei, P., Agbogah, K., Bannerman, P., Sotor, M., Yamoah, K. K., Bilisini, W. B., 2018. Status of the small pelagic stocks in Ghana and recommendations to achieve sustainable fishing 2017. Scientific and Technical Working Group. USAID/Ghana Sustainable Fisheries Management Project (SFMP). Coastal Resources Center, Graduate School of Oceanography, University of Rhode Island. GH2014_SCI042_CRC 22 pp.

Larson WA, Seeb LW, Everett MV et al. (2014) Genotyping by sequencing resolves shallow population structure to inform conservation of Chi- nook salmon (Oncorhynchus tshawytscha). Evolutionary Applications, 7, 355–369.

Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 2009;25:1754–1760. doi: 10.1093/bioinformatics/btp324.

Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics. 2010;26:589–595. doi: 10.1093/bioinformatics/btp698.

Li W, Godzik A. 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22:1658-1659

Longhurst, A.R. and Pauly, D., 1987. Ecology of Tropical Ocean. Academic Press Inc. Harcourt Brace Jovaanovich, Publishers, San Diego. 407p.

Meirmans, P. G. & Van Tienderen, P. H., 2004. Genotype and Genodive: Two Programs for the Analysis of Genetic Diversity of Asexual Organisms. Mol. Ecol. Notes 4, 792–794.

Narum SR, Buerkle CA, Davey JW, Miller MR, Hohenlohe PA, 2013. Genotyping-by Sequencing in ecological and conservation genomics. Molecular Ecology, 22, 2841–2847.

Nielson, J. 1995. Evolution and the aquatic ecosystem: defining unique units in population conservation. American Fisheries Society, Symposium 17, Bethesda, Maryland.

Palumbi, S.R., Baker, C.S., 1994. Contrasting population structure from nuclear and mtDNA of humpback whales. Mol Biol Evol 11:426–435.

Puritz, J. B., Hollenbeck, C.M., Gold, JR., (2014) dDocent: a RADseq, variant-calling pipeline designed for population genomics of non-model organisms. *PeerJ* 2: e431.

Puritz, J.B., Gold, J.R. and Portnoy, D.S., 2016. Fine-scale partitioning of genomic variation among recruits in an exploited fishery: Causes and consequences. *Scientific reports*, *6*, p.36095.

Pukk, L., Kuparinen, A., Gross, R., Vasem.gi, A., 2013. Genetic and life-history changes associated with fisheries-induced population collapse. Evolutionary Applications, 6 (5), 749−760.

Rahimi, P., S, R. G., P, G. M., Sh, J., & Rahnema, M. (2016). Population genetic structure of the white sardine , Sardinella albella , in the Persian Gulf and Sea of Oman by analysis of mitochondrial control region, *15*(3), 995–1008.

Rousset, F., 2008. GENEPOP ‘ 007: a complete re-implementation of the GENEPOP software for Windows and Linux. Molecular Ecology Resources 8:103–106.

Roldan, M.I., Perrotta, R.G., C0rtey, M. and Pla, C., 2000. Molecular and morphologic approaches to discrimination of variability patterns in chub mackerel *Scomber japonicus*. Journal of Experimental Marine Biology and Ecology, 253, 63-74.

Sinclair, M., 1998. Marine populations: an assay on population regulation and speciation. Washington Sea grant, University of Washington Press, Seattle.

Sukumaran,S., Sebastian, W., Gopalakrishnan, A. Population genetic structure of Indian oil sardine, *Sardinella longiceps* along Indian coast Gene, 576 (2016), pp. 372-378.

Sukumaran, S., Sebastian, W., & Gopalakrishnan, A., 2017. Genetic population structure and historic demography of Indian mackerel , Rastrelliger kanagurta from Indian Peninsular waters. Fisheries Research, 191, 1–9.

Svedang, H., Righton,D., Jonsson, P. Migratory behaviour of Atlantic cod *Gadus morhua*: natal homing is the prime stock-separating mechanism Mar. Ecol. Prog. Ser., 345 (2007), pp. 1-2.

Tacon, A.G.J., 2004. Use of fish meal and fish oil in aquaculture: a global perspective.

Taylor, E. B., Redenbach, Z., Costello, A. B., Pollard, S. J. and Pacas C. J. 2001. Nested Analysis of genetic variation in northwestern North American char, Dolly Varden (*Salvelinus malma*) and bull trout (*Salvelinus confluentus*). Canadian Journal of Fisheries and Aquatic Sciences 58:406–420.
Aquatic Resources, Culture and Development, 1: 3-14.

Teacher, A.G.F., Andre, C., Jonsson, P.R. Merila, J.Oceanographic connectivity and environmental correlates of genetic structuring in Atlantic herring in the Baltic Sea Evol. Appl., 6 (2013), pp. 549-567.

Tudela, S., Garcya - Marynn, J.L. and Pla, C., 1999. Genetic structure of the European anchovy, Engraulis encrasicolus, in the north-west Mediterranean. Journal of Experimental Marine Biology and Ecology, 234, 95-109.

Ward, R.D., Woodmark, M., Skibinski, D., 1994. A comparison of genetic diversity levels in marine, freshwater and anadromous fishes. *Journal of Fish Biology* 44, 213-232.

Wang, L., Liu, S., Zhuang, Z., Guo, L., Meng, Z., Lin, H., 2013. Population genetic studies revealed local adaptation in a high gene-flow marine fish, the small yellow croaker (Larimichthys polyactis) PLoS One, 8 (12), p. e83493.

Weir, B.S., Cockerham, C.C. (1984) Estimating F-statistics for the analysis of population structure. Evolution 38, 358-1370.

White, M. M., Kassler, T. W., Philipp, D.P., and Schell, S.A., 2005. A genetic assessment of Ohio River walleyes. Transactions of the American Fisheries Society 134:661–675.

Whitehead, P.J.P. Clupeid fishes of the world. An annotated and illustrated catalogue of the herrings, sardines, pilchards, sprats, anchovies and wolf-herrings. Part 1 – Chirocentridae, Clupeidae and Pristigasteridae. FAO Fish. Synop, 125 (7) (1985), p. 303.

Wilson, R. R. Jr., et Alberdi, P. D. Jr., 1991. An electrophoretic study of Spanish sardine suggests a single predominant species in the eastern Gulf of Mexico, *Sardinella aurita*. Can. J. Fish. Aquat. Sci., Vol. 48: 792-798.

Willing EM, Dreyer C, van Oosterhout C (2012) Estimates of genetic differentiation measured by Fst do not necessarily require large sample sizes when using many SNP markers. PLoS ONE 7:1–7.

Willis, S. C., Hollenbeck, C. M., Puritz, J. B., Gold, J. R. and Portnoy, D. S. 2017, Haplotyping RAD loci: an efficient method to filter paralogs and account for physical linkage. Mol Ecol Resour, 17: 955–965. doi:10.1111/1755-0998.12647.

Zizah S., Hilmi, K., Laarissi, J. and Makaoui, A. 2001. Impact des conditions du milieu marin sur les variations de l'abondance de la sardine dans la zone centrale (Safi-Cap Boujdor) du Maroc. Bulletin de l'Institut Scientifique, Rabat, Section Sciences de la Vie, 2001, No. 23: 37-45.